

VU Research Portal

Spelling Dutch doublets: Children's acquisition of a morphological spelling pattern

Notenboom, A.; Reitsma, P.

published in

Scientific Studies of Reading
2007

DOI (link to publisher)

[10.1080/10888430709336556](https://doi.org/10.1080/10888430709336556)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Notenboom, A., & Reitsma, P. (2007). Spelling Dutch doublets: Children's acquisition of a morphological spelling pattern. *Scientific Studies of Reading*, 11, 133-150. <https://doi.org/10.1080/10888430709336556>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

This article was downloaded by: [Vrije Universiteit, Library]

On: 20 April 2011

Access details: Access Details: [subscription number 907218003]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Scientific Studies of Reading

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t775653700>

Spelling Dutch Doublets: Children's Learning of a Phonological and Morphological Spelling Rule

Annelise Notenboom^a; Pieter Reitsma^a

^a PI Research—Vrije Universiteit Amsterdam,

To cite this Article Notenboom, Annelise and Reitsma, Pieter(2007) 'Spelling Dutch Doublets: Children's Learning of a Phonological and Morphological Spelling Rule', Scientific Studies of Reading, 11: 2, 133 — 150

To link to this Article: DOI: 10.1080/10888430709336556

URL: <http://dx.doi.org/10.1080/10888430709336556>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Spelling Dutch Doublets: Children's Learning of a Phonological and Morphological Spelling Rule

Annelise Notenboom and Pieter Reitsma
PI Research—Vrije Universiteit Amsterdam

This study addresses the question of why spellings determined by morphology are relatively hard to acquire by presenting a latent class model of children's acquisition of a doublet of consonants in the spelling of Dutch verbs. This spelling pattern can be determined either by a phonological rule (after a short vowel, a doublet is spelled) or a morphological rule (doublets are spelled in past-tense forms). The results show that the youngest group of spellers identified by latent class analysis predominantly used an alphabetic strategy. They did not spell doublets at all. The latent class model further shows that the acquisition of phonologically determined spellings occurred at a lower average age than the acquisition of morphologically determined spellings. The latter led to overgeneralizations, and a U-shaped developmental pattern was found as a consequence of these overgeneralizations. Children overgeneralized doublets for different reasons. At younger ages, overgeneralizations of doublets occurred because children treated the doublet as a phonological alternative to the singleton, whereas at older ages, overgeneralizations of doublets were confined to homophones, indicating lexical effects.

Some words are notoriously hard to spell. For example, even competent English spellers may doubt the correct spelling of the genitive apostrophe, as in *pupils' entrance*, given alternatives such as *pupil's entrance* and *pupils entrance*. Spellings that are determined by grammatical rules, such as the genitive apostrophe, are prone to error. Why are the elements of the writing system that are related to grammar so difficult? Why are these elements of the spelling system so hard to acquire? The present article addresses these questions.

Fundamental to the writing systems of alphabetic languages is the alphabetic principle, which states that phonemes (single distinctive sounds) are mapped onto graphemes (letters or letter combinations). Thus, the word *tree* consists of three phonemes, /t/, /r/, and /i/. Now the problem arises that in most alphabetic orthographies, the correspondences between phonemes and graphemes are often not one-to-one, but one-to-many or many-to-one. For example, in English, the vowel /o/ is spelled in at least 13 different ways, as in *dole*, *droll*, *bowl*, *coal*, *toe*, *folk*, *soul*, *owe*, *sew*, *dough*, *mauve*, *brooch*, and *yeoman* (Barry, 1994). The preservation of etymology is one important reason for the inconsistency between phonemes and graphemes; for example, the *k* in *knot* was once pronounced.

Sometimes phoneme–grapheme inconsistencies can be resolved when the position of phonemes is considered. For example, in English, the letter combination *ck* as in *clock* does not occur at the beginning of a word. Taking the following or preceding letters into account can also disambiguate phoneme–grapheme relations (Kessler & Treiman, 2001; Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995).

Phoneme–grapheme inconsistencies may also reflect consistencies at the morphological level of the language, where morphemes are defined as combinations of phonemes that have a semantic or grammatical function and cannot be further subdivided into smaller parts. For example, the morpheme *-ed* marks the past tense of regular English verbs, although it is differently pronounced in *walked* and *ended*.

During the spelling process, a competent speller integrates different types of knowledge, such as phoneme–grapheme correspondences, word context, letter position, morphology, or word-specific idiosyncrasies, derived from different linguistic levels. The question arises as to how people acquire these different types of knowledge. According to developmental theories of spelling, types of spelling knowledge develop at different times. Stage theories assume that children use different strategies along the developmental continuum, and that there is a qualitative shift when children move from one stage to the next (Bear & Templeton, 1998; Ehri, 1992; Henderson, 1992). There is less agreement on the identification and characterization of the different stages, although most models include a prephonological, a phonological, and an orthographic stage, marking the acquisition of knowledge of phoneme–grapheme correspondences and the use of orthographic knowledge such as word context, letter position, and morphology. However, stage models have not gone without criticism. According to Overlapping Waves Theory, children use a variety of strategies or sources of knowledge, which compete with each other over prolonged periods of time (Siegler, 2000). Changes in strategy use may occur, but these are gradual and have varying rates. The model of overlapping waves could also account for spelling (Rittle-Johnson & Siegler, 1999).

There is, however, considerable evidence to support the hypothesis derived from stage theory that children learn the phonological principle of spelling long before they learn to use the conventional spellings for morphemes (Nunes, Bryant, & Bindman, 1997a). At initial phases of spelling development, children do not

mark silent morphemes or spell nonsilent morphemes phonologically. Then, a phase of overgeneralization of the alternative spelling or spellings follows, after which these spellings are confined to the right group of words. So, for example, when learning the spelling of past-tense regular verbs in English, children initially spell words phonologically: *kist*, *slept*, and *soft* for *kissed*, *slept*, and *soft*, respectively. Then, children learn the past-tense inflection but overgeneralize it to inappropriate words, which results in *kissed*, *sleped*, and *sofed*. Finally, words are correctly spelled (examples are taken from Nunes et al., 1997a). More important, this developmental sequence is found for several other alphabetic languages, including Greek (Chliounaki & Bryant, 2002; Harris & Giannouli, 1999), Danish (Juul, 2005; Juul & Elbro, 2004), English (Bryant, Nunes, & Bindman, 2000; Nunes et al., 1997a, 1997b), French (Totereau, Barrouillet, & Fayol, 1998; Totereau, Thevenin, & Fayol, 1997), Dutch (Assink, 1985; Frisson & Sandra, 2002; Sandra, Frisson, & Daems, 1999), and Brazilian Portuguese (Rego, 1999). Thus, although the level of transparency, the particular morphemes, and the details of the orthography vary vastly for these scripts, the psychological processes involved in learning phonological and morphological structures may be very similar.

An interesting question is why a phase of overgeneralization occurs during the acquisition of morphological spellings. Probably children learn alternative spellings without understanding their morphological basis (Bryant, Nunes, & Aidinis, 1999) and therefore simply treat them as phonological alternatives (Chliounaki & Bryant, 2002). However, frequency effects are found in spellers' productions of homophones, as is true for instance for *marche* and *marchent* in French (Totereau et al., 1998) or *rijd* and *rijdt* in Dutch (Assink, 1985; Frisson & Sandra, 2002). This suggests that choosing between alternative spellings is at least partly lexically determined. These accounts are not incompatible; possibly, there are different causes for overgeneralizations at different points in the developmental course. The main goal of the present study was to further clarify the nature of overgeneralizations along the developmental continuum by studying the acquisition of the spelling of regular past-tense Dutch verbs.

In Dutch spelling of verbs, the morphemes marking a past tense are *-te* or *-de*. These morphemes can be followed by an *n* to indicate plurality. The morphemes are added to the stem of the verb. So for instance, the stem of the verb *werken* (to work) is *werk*; the past-tense forms are *werkte(n)*. It is important to note that the morphemes do not exclusively mark past-tense forms in Dutch. They occur in several verbal forms and in other word categories such as nouns and adjectives. For instance, the string *te* occurs in the adjective *zachte* (soft). The difficulty with past-tense forms is that sometimes the stem of the verb ends with a *t* or *d*. For instance, the stem of the verb *wachten* (to wait) is *wacht*; and consequently, the past-tense form is *wachtte(n)*, resulting in a doublet of *t*'s. There is no phonological difference between singletons and doublets here. The plural past-tense form *wachtten* is homophonous to the infinitive and plural present tense *wachten*. The

present study focuses on the acquisition of the past-tense forms that include doublets of *t*'s and *d*'s. Yet doublets are not the exclusive hallmark of past-tense forms. In Dutch, doublets follow short vowels at syllable boundaries (compare the English words *swimming* and *funny*). Consequently, doublets of all consonants, including *t*'s and *d*'s, can occur in all word categories, including verbal forms. For instance, the participle *uitgeputte* (exhausted) is not a past-tense form but written with a doublet because the vowel *u* is short. Thus, tracking children's learning of the spelling of doublets will give insight in the acquisition of a spelling pattern that can be determined by a phonological rule (that doublets follow a short vowel) or by a morphological rule (that in past-tense forms, the morpheme *-te* or *-de* is added to the stem of the verb).

In the Netherlands, the general rule of the spelling of doublets following short vowels is taught to children at the second or third grade, whereas it is not until the fifth grade that children specifically are taught the rules for the spelling of verbal forms. Based on the findings of previous research discussed above, we hypothesize that at the onset of development, children spell all verbal forms phonologically (i.e., with singletons). Then, children learn to spell doublets following short vowels (Landerl & Reitsma, 2005). No overgeneralizations are hypothesized to occur as a consequence of learning this phonological rule. However, if children start to use the morphological rule and write doublets in past-tense forms correctly, they will overgeneralize them to other verbal forms. If children overgeneralize the doublet because they treat it as a phonological alternative, they will do so in all verbal forms. They will write *werkten* instead of *wachten* as often as *wachtten* instead of *wachten*, although the last three are orthographically legal while the first is not. On the other hand, if overgeneralizations are lexically determined, the misspelling *werkten* will have a much lower rate than *wachtten*—the more so if the latter has a higher frequency rate than *wachten*. Thus, studying the spellings of phonologically determined or morphologically determined doublets may give insight into the incidence and causes of overgeneralizations.

Usually, the developmental process is modeled by analyzing the response patterns of children of different age groups, including their types of errors. Recent studies show that latent class analysis (LCA) can be a useful technique in modeling cognitive development (Boom, Hoijtink, & Kunnen, 2001; Hoijtink & Notenboom, 2004; Jansen & Van der Maas, 1997; Raijmakers, Jansen, & Van der Maas, 2004). Given, for example, the responses on a spelling test, the aim is to infer which strategies children use. Responses on a spelling test are directly observable, unlike the use of a strategy; responses and underlying strategies can be theorized to correspond to the two types of variables of the latent class model: observed categorical variables and an unobserved (e.g., latent) categorical variable. Basically, LCA classifies children in homogeneous groups (latent classes) given their observed responses. The categories of the latent variable can be viewed as corresponding to the strategies children use, which help explain the associations between the observed responses. The

use of (exploratory) LCA has two main advantages in the context of developmental research. First, strategies need not be defined beforehand. It is possible that a greater number or variety of strategies would be detected than theoretically expected. Second, LCA gives a clear indication of the fit of the model to the data by selecting and interpreting the most parsimonious model.

This study presents a latent class model on the acquisition of the spelling of Dutch doublets. Latent classes are hypothesized to be formed corresponding to children's strategies described above. Thus, we expect to find one latent class consisting of alphabetic spellers, who write all words with singletons. Children who know how to spell doublets after short vowels but not in past-tense forms are hypothesized to form a second latent class. If children learn to spell doublets in past-tense forms, they may overgeneralize them to other verbal categories. Children who do so are hypothesized to form the third latent class. Finally, we expect to find a fourth latent class of children who confine the incorrect spelling of doublets to the homophones. Because these incorrect spellings are lexically determined, frequency effects are to be expected.

METHOD

Participants

Participants were 458 students from the third to the sixth grade of elementary education and the two adjacent grades of high school. Two schools of elementary education and two high schools from the Amsterdam area participated in the experiment. In the Netherlands, the grades of high school are separated in three levels; classes from all three levels participated. A summary of the number of students per grade is given in Table 1. The elementary schools reflected the multiethnic population of Amsterdam, whereas the students of the high schools were predominantly nonimmigrant.

TABLE 1
Number and Average Age of Students Per Grade

<i>Grade</i>	<i>No. of Students</i>	<i>Average Age (Years; Months)</i>
3 Elementary education	42	9;3
4 Elementary education	50	10;5
5 Elementary education	52	11;6
6 Elementary education	36	12;5
1 High school	160	12;9
2 High school	118	13;9
Total	458	
Average		12;4

Materials

A spelling test of Dutch verbal forms was administered to all participants. The items of the spelling test were 32 verbal forms that contained the strings *-te* or *-de*. There were seven different categories of verbal forms included in the test. These categories are summarized in Table 2. The first two categories include verbs that are spelled with doublets, but for different reasons. The first is related to the phonological rule that doublets follow a short vowel and the second is related to the morphological rule that the morpheme *-te* or *-de* is added to the stem. The following five categories of verbal forms are all spelled with singletons. It is important to note that phonologically, there is no difference between singletons and doublets. The occurrence of doublets of *t*'s and *d*'s allows study of the acquisition of an alternative morphological spelling and possible overgeneralizations of this spelling across different verbal categories.

The first category of verbal forms in Table 2 includes doublets of *t*'s and *d*'s marking short vowels at syllable boundaries. This phonological spelling rule is fairly fundamental in Dutch and surpasses all specifications of the spelling of verbal forms mentioned below. The second category of Table 2 consists of verbal forms of which the stem ends with a *t* or *d*. These verbal forms get a doublet in the past tense. Clearly, the doublets of Category 2 are morphologically determined.

TABLE 2
Categories of Verbal Forms Included in the Test

Category: Average Frequency of Items	No. of Items	Examples			
		Verb	Translation	Stem	Verbal Form
1 Doublets following a short vowel [65]	4	uitputten wedden	to exhaust to bet	uitput wed	uitgeputte wedde
2 Weak past tense verbal forms; stem ends in <i>t</i> or <i>d</i> [495]	8	wachten landen	to wait to land	wacht land	wachtte(n) landde(n)
3 Weak past tense verbal forms [803]	6	werken vieren	to work to celebrate	werk vier	werkte(n) vierde(n)
4 Strong verbs [3,203]	4	vechten houden	to fight to like	vecht houd	vocht(en) hield(en)
5 Participles used as adjectives [101]	2	blussen huilen	to extinguish to cry	blus huil	gebluste huilende
6 Infinitives ^a [1,151]	4	haasten bevrijden	to hurry to liberate	haast bevrijd	haasten bevrijden
7 Participles used as adjectives; stem ends with <i>t</i> or <i>d</i> ^b [108]	4	stranden verbreden	to strand to broaden	strand verbreed	gestrande verbrede

^aCan be homophones to the past-tense plural. ^bCan be homophones to the past-tense singular.

The verbal forms of the third to the seventh categories are all spelled with singletons. The third category includes past-tense forms in which the stem ends with any letter except for a *t* or *d*. Basically, the morpheme sequences *-te(n)* or *-de(n)* are added to the stem, depending on a voiced or voiceless final consonant, whereas the final *n* marks the plural form. In Dutch, there are also strong verbs, belonging to Category 4 in Table 2. Strong verbs are always spelled phonologically (e.g., with a singleton). Participles used as adjectives form Categories 5 and 7 in Table 2. The correct spelling of these participles includes a singleton, no matter whether the stem ends with a *t* or *d* (Category 7) or not (Category 5). Participles can be homophonous to singular past-tense forms. If the stem of the verb ends with a *t* or *d*, there are therefore two legal spellings of these homophones—one including a doublet and the other a singleton. Choosing the correct spelling requires grammatical analysis. The sixth category of Table 2 is infinitives spelled with singletons. If the stem of verbal forms ends with a *t* or *d*, these infinitives are homophonous to the plural forms of the past tense. Again, there are two legal spellings of these homophones—one including a singleton, the other a doublet—and again, the correct spelling reflects grammatical understanding.

All seven categories of verbal forms just discussed were included in the test. The sentences were fairly short: The average number of words was 7.3. Previous research indicates that the complexity of the grammatical structure of the sentence affects the frequency of correct answers (Assink, 1985). In the present study, the grammatical structure of the sentences was simple: The subject of the sentence, or in the case of participles, the relevant substantive, was placed immediately before or after the verbal form, facilitating correct answers. Table 2 presents the number and average frequency of items included in the test that represent the categories of the verbal forms. All items were fairly frequent to highly frequent. In particular, the strong verbs included in our test were highly frequent, as is generally the case for strong verbs. All infinitives (Category 6) and half of the participles (of Category 7) were homophonous to past-tense forms. The other half of the participles differed from homophonic forms in the addition of affixes only. The orthographic word forms of items in Category 6 and 7 included are equally as frequent or more frequent than their homophones (according to the 42 million CELEX count of Baayen, Piepenbrock, & Van Rijn, 1993). For example, the verbal form *haasten* included in the test has a frequency count of 129, whereas its homophone *haastten* has a frequency count of 108. Given the equivalence of frequencies, it is to be expected that participants would tend to select the correct orthographic form without regard to frequency.

Procedure

The spelling test that was administered to all participants consisted of 32 printed sentences that were read aloud by the experimenter or teacher. The 32 items of the

test were verbal forms that were commonly known to children of elementary school age. Participants followed the printed version of the sentence on their response sheet. After the verb form was reiterated, participants filled in the blank space. Responses were scored as either correct or incorrect, focusing on whether the correct number of *t*'s or *d*'s was spelled and ignoring possible spelling errors in other parts of the words.

Analyses

The data set was analyzed by LCA. An introduction to this statistical method can be found in Clogg (1995), Hagenaaers (1993), McCutcheon (1987), and Rost and Langeheine (1997). A latent class model defines two types of categorical variables. First, there are manifest, observed variables (here, the items of the spelling test). Second, there is a latent, unobserved variable. The categories of the latent variable are called *latent classes*. Exploratory LCA reduces the set of observed variables to a single latent variable. (Models with more than one latent variable are not considered here). The associations between the observed variables are explained by the latent variable with two or more latent classes. Basically, students with similar responses on the observed variables are assigned to the same latent class.

A latent class model is typically interpreted by means of the class weights and response probabilities. The class weights are the proportions of students that belong to particular latent classes. The response probabilities indicate the probability that a student in a latent class will have a particular score (here, correct or incorrect) on the observed items. Response probabilities represent the degree of association between the observed variables and the latent classes and are therefore used to interpret the nature of the latent classes. Response probabilities are the same for all students within a latent class.

Estimation of the model was based on Bayesian computational methods. Gibbs sampling was used to estimate the parameters of the latent class model (for a formal account, see Hoijtink, 2001; Hoijtink & Notenboom, 2004). The latent classes were defined using an algorithm based on the principles described by Richardson and Green (1997). During the first iteration, the algorithm splits the whole sample into two latent classes. During subsequent iterations, the algorithm seeks to split the latent classes further, and children are allowed to move from one latent class to any other. If none of the latent classes can be split any more, the maximum number of latent classes has been reached. Requesting an additional latent class leads to an empty class (i.e., one without a member).

A sequence of latent class models was estimated, each with a different number of classes. The selection of the best model was determined by using two fit measures, the pseudo-likelihood ratio test and the $-2 \log$ marginal likelihood (Hoijtink, 2001). The pseudo-likelihood ratio test is an absolute fit measure and compares for each pair of items the response probabilities that are observed with the probabili-

ties predicted from the latent class model. The pseudo-likelihood ratio test is evaluated using a p value. A large p value indicates that the predicted values are within the limits of chance variation from the observed values for each pair of items. The marginal likelihood is a relative fit measure, and a Bayesian counterpart of information indices like (corrected) Akaike's Information Criterion (Kass & Raftery, 1995). An information criterion is a measure of the distance between the model at hand and the true model. It takes into account both the statistical goodness of fit and the number of parameters that have to be estimated to achieve that degree of fit. Lower values indicate a more parsimonious fit to the data. The selection of the best model is based on its goodness of fit, its parsimony, and its interpretability. The software used to estimate the present latent class model is based on the study of Hoijtink and Notenboom (2004).

RESULTS

Cronbach's alpha coefficient for internal consistency of the test was .58, which is relatively low, probably because different categories of verbal forms were included. A first presentation of the results is given by computing the average correct scores for each grade level and for the seven categories of verbal forms. The results are graphically presented in Figure 1. On the horizontal axis, the six grades are given. Each verbal category is represented by a distinct line. Clearly, children at the third and fourth grades of elementary school had difficulties in spelling doublets, whether they are phonologically determined (Category 1) or morphologically determined (Category 2). However, they had perfect scores on all remaining categories that are spelled with singletons (Categories 3–7). Children at higher grades got increasingly better at spelling phonologically determined doublets, and to a lesser extent, at morphologically determined doublets. However, learning doublets appeared to have a detrimental effect on the correct spelling of singletons. From Grade 4 onward, children overgeneralized the doublet to verbal forms that are correctly spelled with singletons. For Categories 3 to 5, a U-shaped developmental pattern was found; intrusions of the doublet faded out at higher grades. However, intrusions of doublets remained for the infinitives (Category 6) and for participles used as adjectives in which the stem ends with *t* or *d* (Category 7) at higher grades.

A latent class model was fitted to categorize children in homogeneous groups. Given types of response patterns, strategies can be inferred. A sequence of models with an increasing number of latent classes was estimated. The values of the pseudo likelihood ratio statistic and the -2 log marginal likelihood of these models are summarized in Table 3. The model with seven latent classes had one empty class (i.e., one without a member and was therefore not considered further). Models with three or more latent classes had p values equal to or greater than .01, indicating that these models had a fairly adequate fit. The model with five latent

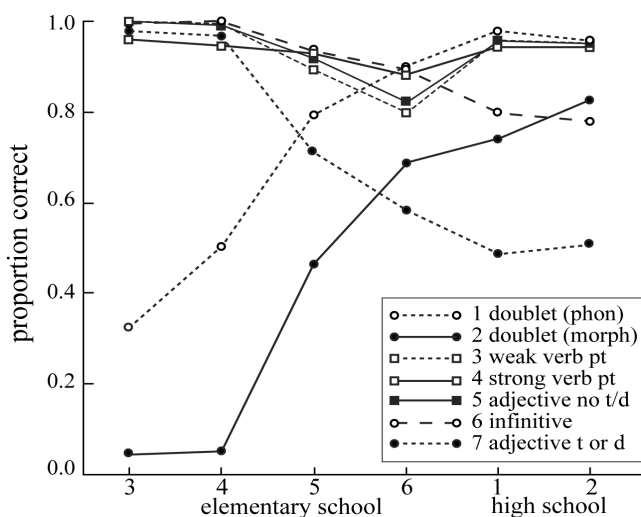


FIGURE 1 Average proportion correct score per grade level (Grades 3–6 of elementary school, and Grade 1, 2 of high school) and category (seven categories of verbal forms; see also Table 2). *Note:* Category 1 = Doublets following short vowels. Category 2 = Doublets are required because the stem ends with a *t* or *d* and the morpheme *-te* or *-de*, indicating the past tense (pt), is added to the stem. Category 3 = Singletons are required in past-tense forms because the stem does not end with a *t* or *d*. Category 4 = Strong verbs are spelled with singletons. Category 5 = Participles used as adjectives are spelled with singletons. Category 6 = Infinitives are spelled with singletons; they can be homophonous to plural past-tense forms, spelled with doublets. Category 7 = Participles used as adjectives are spelled with singletons—the stem ends with *t* or *d*; they can be homophonous to singular past-tense forms spelled with doublets.

TABLE 3
Fit Measures of Latent Class Models

No. of Latent Classes	–2 Log Marginal Likelihood	<i>p</i> Value Pseudo Likelihood Ratio Test
2	9416	.00
3	9066	.01
4	8887	.02
5	8873	.02
6	8792	.01

classes was selected and interpreted because it had a relatively high *p* value and a low marginal likelihood, but alongside, remarks will be made on the models with three to six latent classes because the fit measures were very close.

The five latent classes differed considerably in size. Table 4 presents the number of children that were assigned to the latent classes and the average age of the children. Latent Class 4 included nearly half of the sample, whereas the other four latent classes were small to intermediate in size. The latent classes are numbered in terms of average age. To interpret the results, average response probabilities were computed for each of the seven categories of verbal forms and for each latent class. The results are presented graphically in Figure 2. Each latent class is signified by a distinct line. The vertical axis represents the average proportion of correct answers. On the horizontal axis, the categories of verbal forms are given.

TABLE 4
Number and Average Age of Students Per Latent Class

<i>Latent Class</i>	<i>No. of Students</i>	<i>Average Age (Years; Months)</i>
1	117	10;7
2	63	12;8
3	17	12;8
4	207	13;2
5	54	13;4

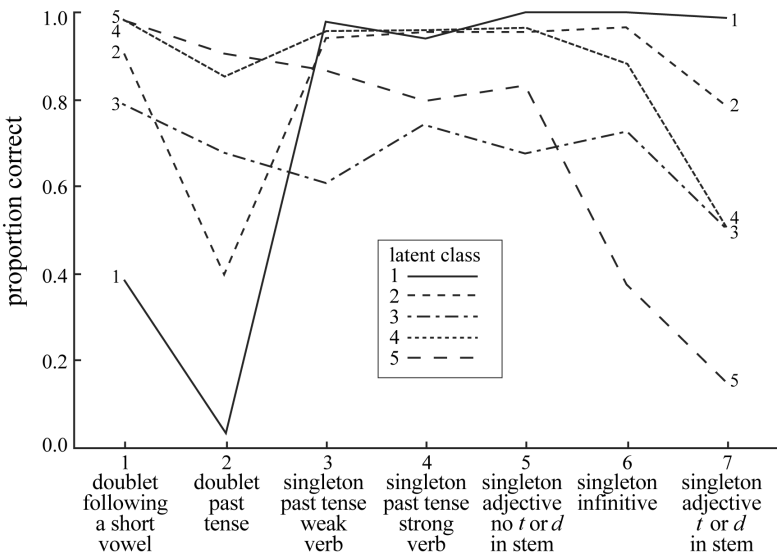


FIGURE 2 Average proportion correct scores per category and latent class.

Figure 2 shows that children of the first latent class, who were the youngest on average, had a poor command of spelling doublets, in particular regarding past-tense forms (Category 2). However, they spelled verbal forms of other categories perfectly. Hence, children of the first latent class can be characterized as alphabetic spellers. The following two latent classes (2 and 3) comprised children of an average age of 12 years 8 months. Children at the second latent class spelled doublets correctly if the doublets follow short vowels (Category 1) but still had difficulties with past-tense forms that include doublets (Category 2). They spelled verbal forms of all other verbal categories nearly correctly. The third latent class was small in number, consisting of 17 children. They spelled the verbal forms involving doublets fairly correctly, but problems arose for the five categories of verbal forms that must be spelled with singletons. A diffuse pattern of overgeneralizations of doublets to all verbal categories was found. In particular, the past-tense forms that are spelled with singletons (Category 3) had low scores, compared to the other latent classes. The occurrence of overgeneralizations at similar rates in all spelling form categories indicates that students at the third latent class treated the doublet as a phonological alternative to the singleton.

The fourth and fifth latent class included the oldest children, who were about 13 years old. The fourth latent class is relatively large, as it includes half of the sample. Children of this latent class spelled all categories properly except for the participles that are used as adjectives in which the stem ends with a *t* or *d* (Category 7). Intrusions of doublets were characteristic for the responses of children at Latent Class 5, in particular with respect to the infinitives (Category 6) and the participles in which the stem ends with a *t* or *d* (Category 7). These categories are homophonous to past-tense forms that are spelled with doublets. Probably, these overgeneralizations are lexically determined. However, the low average percentage of correct scores is remarkable, given the fact that the correct spellings have an equal or higher frequency count than the incorrect spellings.

The model with three latent classes differs from the one interpreted here in the following respects. Latent Classes 1 and 2 are taken together, and Latent Class 3 is not distinguished. Thus, according to this model, there is one class of children who have a poor command of doublets and two latent classes of children who have a good command of doublets but who overgeneralize in particular to the homophonic categories. The model with four latent classes splits the latent class of children with poor command of doublets similarly to the model with five latent classes previously described but does not identify Latent Class 3. The model with six latent classes is also comparable to the model with five latent classes. The additional sixth latent class is parallel to the third latent class depicted in Figure 2, although the average correct response of all categories is higher (around .80). The sixth latent class provides additional evidence for a diffuse pattern of overgeneralizations, although overall error rates are fairly low.

DISCUSSION

Why are morphologically determined spellings such as the Dutch doublet so difficult, and why are they so hard to acquire? The present study addresses these questions by presenting a latent class model on the acquisition of the doublet in Dutch, a spelling category that can be determined by either a phonological or a morphological rule. At initial phases of development, children predominantly use an alphabetic strategy. The youngest group of spellers (about 10 years of age) identified by LCA spelled words according to phoneme-grapheme correspondence rules: These children were mostly unaware of the grammatical and morphological categories reflected in the writing system and did not spell any doublets. Results from other studies support the finding that learning the alphabetic principle is very fundamental in learning to spell (Cassar, Treiman, Moats, Pollo, & Kessler, 2005; Ehri, 1998; Notenboom & Reitsma, 2003).

The second latent class consisted of children with an average age of 12 years 8 months. They knew how to spell phonological doublets (following short vowels), but they had poor command of morphologically determined doublets (occurring in past-tense forms). Clearly, the phonological spellings were acquired earlier than morphological spellings. No substantial rates of overgeneralizations of doublets to other verbal forms were shown. Children in Latent Class 3 were about the same age as children in Latent Class 2. They spelled phonological and morphological doublets correctly but overgeneralized them to all other verbal forms. Given this diffuse pattern of overgeneralizations, children in Latent Class 3 seemed to use the doublet as a phonological alternative to the singleton. They did not have much understanding of its morphological basis. However, evidence for this type of strategy was not overly robust. The number of children in Latent Class 3 was fairly small, and these children were not discerned as a distinct group in the models with three or four latent classes.

Learning morphological doublets resulted in a higher rate of overgeneralizations than learning phonological doublets did, given the difference between Latent Class 2 (characterized by the acquisition of phonological doublets but few overgeneralizations) and Latent Class 3 (characterized by acquisition of both phonologically and morphologically determined doublets and overgeneralizations). The occurrence of overgeneralizations thus shows the beginnings of children's understanding of the written representations of morphological aspects of the language.

Children of the fourth and fifth latent class were good at spelling phonological and morphological doublets. They were the oldest of the sample (i.e., around 13 years old). In particular, children in Latent Class 5 overgeneralized the doublet, above all to those categories in which a homophone exists, but to a lesser extent also to the other categories). Latent Class 4, which contained almost half of the sample, was similar to Latent Class 5, except that children in Latent Class 4

overgeneralized the doublet to only one category in which a homophone exists (the participles that are used as adjectives). Thus, at higher ages, the overgeneralizations had mainly lexical origins.

Clearly, learning the doublet as an alternative morphological spelling had a negative effect on the correct spelling of regular words. For example, children at the fourth and fifth grade of elementary education have a 99% correct score for infinitives spelled with singletons (Category 6), whereas children at the first and second grade of high school have only a 80% correct score. It is likely that this percentage increases as children get older. However, while spelling homophones, older participants remain sensitive to frequency relations. Probably not even adults will reach the 99% correct score, because they will also be inclined to spell the most frequent form. The U-shaped, falling-rising developmental pattern has been reported most often in the context of spelling acquisition (Bryant et al., 1999; Juul & Elbro, 2004) but also occurs for other aspects of language development. For example, when young children learn the past-tense inflection in English, they pass from a period of correct performance through a phase of regularization of irregular verbs (i.e., *finded* or *comed*; Marcus et al., 1992). Marcus et al.'s explanation of this phenomenon is that children's memory of irregulars is not strong enough to block the rule of past-tense inflection. The more frequent a wordform is, the stronger a particular memory trace will be.

The participles of Category 7 and to a lesser extent the infinitives of Category 6 showed high error rates. Paradoxically, participles are orthographically regular and notoriously difficult. One source of difficulty is probably that people think of participles as verbal forms, whereas they are spelled as adjectives. In particular, the spelling of past participles may be very confusing. In addition, the context of an official spelling test can be misleading, because children may think that the more difficult form is being asked for. To test this, possible differences in results between free writing and a spelling test need to be examined.

Other linguistic aspects of the test items may also have had an influence on error rates. Data of Latent Class 2 suggest that doublets show up relatively often in words with vowels at syllable boundaries. Thus, children in Latent Class 2 spell past-tense forms having doublets and vowels at syllable boundaries, such as *raadde* (guessed), more often correctly (62%) than equivalent past-tense forms having doublets and consonants at syllable boundaries, such as *wachtten* (waited; 33% correct). Moreover, participles with vowels at syllable boundaries show more frequent intrusions of doublets (32%), such as *verbreedde* instead of *verbrede* (broadened) than participles with consonants at syllable boundaries (13%), such as *gestrande* (stranded). However, more verbal forms are needed to confirm this effect.

Basically, the results of the latent class model confirmed our hypotheses. At initial phases, children spelled alphabetically and did not use doublets at all. Phonologically determined doublets were acquired earlier than morphologically determined doublets, partly reflecting the order in which the forms are taught in

elementary school. Children who incorrectly generalized the doublet to all verbal forms were relatively young and small in number. Thus, there is some evidence that children treat the doublet as a phonological alternative to the singleton. Intrusions of the alternative spelling appeared in all words at the same rate. As children got older (and presumably lexical memory grew stronger), impossible spellings of words gradually disappeared. Older children incorrectly generalized the doublet mainly to homophonic forms, indicating that their misspellings were lexically determined.

The results demonstrate that LCA is a statistically reliable way to assess strategies children use, given their dichotomous responses on a spelling test. Compared to averaging responses per grade, the latent class model shows the diversity of the response patterns in the data, from which children's strategies can be inferred. It is possible that LCA detects unanticipated strategies because it does not require previously designed criteria. Furthermore, the fit of the model is tested on the whole data set, such that the most parsimonious model can be selected based on statistical criteria.

It is tempting to think of the latent class model as a developmental one. In particular, it is an interesting question whether spellers usually go through Latent Class 3 or if only some go through such a phase in which doublets are a phonological variant. However, the developmental path of individual children can be different from the model just outlined. Furthermore, it may be important to know how long developmental phases will take. Longitudinal data are needed to answer these questions and to trace the developmental course in detail. Nevertheless, the results from the latent analysis partly support a phase model of the acquisition of morphological spelling knowledge. Latent Classes 1 to 4 identify qualitatively distinct answer patterns, and hence reveal different strategies for solving spelling problems. Thus, the initial use of an alphabetic strategy, second, learning the morphological spelling without understanding its morphological basis, and third, restricting the morphological spelling to the appropriate categories are probably valid developmental categories. However, the difference between Latent Classes 4 and 5 seems to be more quantitative in nature. The students in Latent Classes 4 and 5 probably used the same strategy, but students in Latent Class 4 just performed better. A continuous growth model may be more appropriate here than a latent class model. A mixed Rasch model can be used to model a mixture of qualitatively and quantitatively different types of responses. This model would allow us to represent quantitative differences among students within a latent class, whereas differences between latent classes are qualitative (Rost & Langeheine, 1997). The development of spelling ability may consist of a series of phases, characterized by qualitatively distinct strategies to solve spelling problems, whereas within a phase, children get increasingly better at performing a particular strategy.

The findings of the present study have several implications for educational practice. First of all, teaching children the spellings of irregular words should not begin until children master the basics of the alphabetic principle. In general, acqui-

sition of orthographic knowledge will be more easily achieved with the alphabetic foundation already in place (Ehri, 1998). Then, children have to become aware of alternative, morphologically based spellings. Effective instruction should include explanation of the relevant grammatical categories and clarification as to which spelling relates to which grammatical category. Such instruction may further increase children's morphological awareness. Probably, one of the major hurdles the speller has to take is realizing that morphological spellings are difficult. Reliance on phoneme–grapheme consistencies or lexical knowledge is often of little use or may even interfere with correct morphological spellings. A possible flaw of weak spellers is that they do not take any syntactic information into account but rather are inclined to base their spelling decisions on nonsyntactical and hence (in the case of morphological spellings) irrelevant information, such as sound structure or frequency. Therefore, realizing that some words can only be correctly spelled through grammatical analysis and not by retrieval from lexical memory probably is a milestone in the acquisition of morphological spelling knowledge.

ACKNOWLEDGMENTS

This research is supported by a grant (NWO 411-21-006) from the Dutch Organization for Scientific Research awarded to Pieter Reitsma.

REFERENCES

- Assink, E. M. H. (1985). Assessing spelling strategies for the orthography of Dutch verbs. *British Journal of Psychology*, 76, 353–363.
- Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Barry, C. (1994). Spelling routes (or roots or rutes). In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27–49). Chichester, UK: Wiley.
- Bear, D. R., & Templeton, S. (1998). Explorations in developmental spelling: Foundations for learning and teaching phonics, spelling, and vocabulary. *The Reading Teacher*, 52, 222–242.
- Boom, J., Hoijtink, H., & Kunnen, S. (2001). Rules in the balance: Classes, strategies or rules for the Balance Scale Task? *Cognitive Development*, 16, 717–735.
- Bryant, P., Nunes, T., & Aidinis, A. (1999). Different morphemes, same spelling problems: Cross-linguistic developmental studies. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 112–133). Cambridge, UK: Cambridge University Press.
- Bryant, P., Nunes, T., & Bindman, M. (2000). The relations between children's linguistic awareness and spelling: The case of the apostrophe. *Reading and Writing: An Interdisciplinary Journal*, 12, 253–376.
- Cassar, M., Treiman, R., Moats, L., Pollo, T. C., & Kessler, B. (2005). How do the spellings of children with dyslexia compare with those of nondyslexic children? *Reading and Writing: An Interdisciplinary Journal*, 18, 27–49.

- Chliounaki, K., & Bryant, P. (2002). Construction and learning to spell. *Cognitive Development*, 17, 1489–1499.
- Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311–359). New York: Plenum.
- Ehri, L. C. (1992). Review and commentary: Stages of spelling development. In S. Templeton & D. R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 307–332). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Frisson, S., & Sandra, D. (2002). Homophonic forms of regularly inflected verbs have their own orthographic representations: A developmental perspective on spelling errors. *Brain and Language*, 81, 545–554.
- Hagenaars, J. A. (1993). *Loglinear models with latent variables* (Sage University Paper series on Quantitative Applications in the Social Sciences, Series No. 07-094). Newbury Park, CA: Sage.
- Harris, M., & Giannouli, V. (1999). Learning to read and spell in Greek: The importance of letter knowledge and morphological awareness. In M. Harris & V. Giannouli (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 51–70). Cambridge, UK: Cambridge University Press.
- Henderson, E. H. (1992). The interface of lexical competence and knowledge of written words. In S. Templeton & D. R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 1–30). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Hojtink, H. (2001). Confirmatory latent class analysis: Model selection using Bayes factors and (pseudo) likelihood ratio statistics. *Multivariate Behavioral Research*, 36, 563–588.
- Hojtink, H., & Notenboom, A. (2004). Model based clustering of large datasets: Tracing the development of spelling ability. *Psychometrika*, 69, 481–498.
- Jansen, B. R. J., & Van der Maas, H. L. J. (1997). Statistical test of the rule assessment methodology by latent class analysis. *Developmental Review*, 17, 321–357.
- Juul, H. (2005). Grammatical awareness and the spelling of inflectional morphemes in Danish. *International Journal of Applied Linguistics*, 15, 87–112.
- Juul, H., & Elbro, C. (2004). The links between grammar and spelling: A cognitive hurdle in deep orthographies? *Reading and Writing: An Interdisciplinary Journal*, 17, 915–942.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.
- Kessler, B., & Treiman, R. (2001). Relationships between letters and sounds in English monosyllables. *Journal of Memory and Language*, 44, 592–617.
- Landerl, K., & Reitsma, P. (2005). Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German. *Journal of Experimental Child Psychology*, 92, 322–344.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. With commentary by Harald Clahsen. *Monographs of the Society for Research in Child Development*, 57(4, Serial No. 228).
- McCutcheon, A. L. (1987). *Latent Class Analysis* (Sage University Paper series on Quantitative Applications in the Social Sciences, 07-064). Newbury Park, CA: Sage.
- Notenboom, A., & Reitsma, P. (2003). Investigating the dimensions of spelling ability. *Educational and Psychological Measurement*, 63, 1039–1059.
- Nunes, T., Bryant, P., & Bindman, M. (1997a). Learning to spell regular and irregular verbs. *Reading and Writing: An interdisciplinary Journal*, 9, 427–449.
- Nunes, T., Bryant, P., & Bindman, M. (1997b). Morphological spelling strategies: Developmental stages and processes. *Developmental Psychology*, 33, 637–649.

- Raijmakers, M. E. J., Jansen, B. R. J., & Van der Maas, H. L. J. (2004). Rules and development in trial classification task performance. *Developmental Review*, 24, 289–321.
- Rego, L. B. R. (1999). Phonological awareness, syntactic awareness and learning to read and spell in Brazilian Portuguese. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 71–88). Cambridge, UK: Cambridge University Press.
- Richardson, S., & Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society, B*, 59, 731–792.
- Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: Variability, choice and change in children's strategy use. *Child Development*, 70, 332–348.
- Rost, J., & Langeheine, R. (1997). A guide through latent structure models for categorical data. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 13–37). Berlin: Waxmann.
- Sandra, D., Frisson, S., & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and Language*, 68, 277–283.
- Siegler, R. S. (2000). The rebirth of children's learning. *Child Development*, 71, 26–35.
- Totureau, C., Barrouillet, P., & Fayol, M. (1998). Overgeneralizations of number inflections in the learning of written French: The case of noun and verb. *British Journal of Developmental Psychology*, 16, 447–464.
- Totureau, C., Thevenin, M. G., & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory and practice across languages* (pp. 97–114). Mahwah, NJ: Erlbaum Lawrence Associates, Inc.
- Treiman, R., Mullenix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107–136.

Manuscript received July 28, 2005

Accepted March 11, 2006